Genetic structure of nomadic Bedouin from Kuwait

T. Mohammad[1,4], **Yali Xue**[2], **M. Evison**[3], and **Chris Tyler-Smith**[2]

[1]University of Sheffield, Division of Genomic Medicine, University of Sheffield, UK

[2]The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK

[3]Department of Anthropology, University of Toronto, Canada

[4]The Leverhulme Centre for Evolutionary Studies, University of Cambridge, UK

Abstract

Bedouin are traditionally nomadic inhabitants of the Persian Gulf who claim descent from two male lineages: Adnani and Qahtani. We have investigated whether or not this tradition is reflected in the current genetic structure of a sample of 153 Bedouin males from six Kuwaiti tribes, including three tribes from each traditional lineage. Volunteers were genotyped using a panel of autosomal and Y-STRs, and Y-SNPs. The samples clustered with their geographical neighbours in both the autosomal and Y-chromosomal analyses, and showed strong evidence of genetic isolation and drift. Whilst there was no evidence of segregation into the two male lineages, other aspects of genetic structure were in accord with tradition.

Keywords

Introduction

Kuwait lies in the Arabian Peninsula at the head of the Persian Gulf. Although it is located near a possible migration route out of Africa that may have been used over 50 000 years ago (Jobling et al., 2004), such early events appear to have had little impact on the current populations of Kuwait and the first known settlement dates only to the 3rd Century BCE when the Ancient Greeks colonized the island of Failaka under Alexander the Great. Historians believe the current settlement was established between the late 16th and 18th Centuries CE (Dickson, 1956). The population of Kuwait is divided into two main groups: townspeople and Bedouin tribal groups (قبائل *Qaba'il*) with affiliations to their "cousins" (ولد عم *wild 'am*) who inhabit other parts of the Arabian Peninsula. The Bedouin are not geographically isolated from each other and are traditionally nomadic, moving from one region to another seeking pasture for their herds to graze and water for survival (Dickson, 1967). The Bedouin conventionally claim descent from two main lineages: from Adnan son of Ishmael—the Adnani lineage—or from Qahtan (or Joktan)—the Qahtani (or Joktani) lineage.

We wished to investigate the genetic structure of the Kuwaiti Bedouin and consider how they compare with neighboring populations, what the genetic consequences of their

traditional lifestyle might be, and whether the social lineages that identify their descent groups are also reflected in their genetic lineages.

The non-recombining part of the Y chromosome was chosen for analysis. Its high levels of population differentiation—a consequence of its low effective population size and, sometimes, of patrilocality—make it generally highly informative regarding genetic structure and differentiation (Jobling and Tyler-Smith, 2003). In addition, its strict paternal inheritance allows comparisons to be made between traditional male-line descent groups and observed male-line genetic structures. It is a single locus, however, and conclusions based on it may not be representative of the rest of the genome. We therefore complemented the Y-chromosome analysis with that of a widely-used panel of autosomal STRs (Collins et al., 2004).

Samples for DNA analysis were collected from volunteers from the Adnani tribes of Al-Aniza, Mutran and Awazim (a Suluba tribe), and the Qahtani tribes of Ajman, Shimar and Murrah. This selection is representative of the most prominent tribes of Adnani and Qahtani lineages of Arabia. Results from the sample were compared with data from other populations derived from the literature.

Materials and Methods

Samples

One hundred and fifty-three distantly related male subjects—known not to share a common ancestor within the last three generations—were collected from six Bedouin tribes. Ethical review was completed by a research ethics committee of The University of Sheffield (UK) and The Kuwait University ethical review board (Kuwait). Informed consent was obtained from all participants. Specimens were collected as buccal swabs. The Bedouin were from the two traditional lineages—Adnani: Mutran (n=30), Aniza (n= 21) and Awazim (n= 37)—and Qahtani: Ajman (n=39), Shimar (n= 21) and Murrah (n= 5; because of the small sample size, Murrah specimens were used only in the autosomal analysis). Samples were collected from four main Kuwaiti Governorates: Awazim and Aniza tribes from the Al-Assima (capital) Governorate (29.4°Lat, 42.2° Long.), Ajman and Murrah from the Al-Ahmadi Governorate (28.7°Lat, 48.3° Long.), the Mutran sample from the Al-Farwaniyah Governorate (29.2°Lat, 47.9° Long.); the Shimar sample from both the Al-Jahrah (29.3°Lat, 47.6° Long.) and Al-Ahmadi Governorates. Locations are shown in Supplementary Figure 1. Kuwaiti Bedouin do not marry outside their own tribe and samples collected in the capital are not admixed. DNA was extracted from buccal cells using the BuccalAmp DNA Extraction Kits and protocol (Epicentre, Madison, WI, USA).

Autosomal STR Typing

Amplification with the Identifiler™ multiplex was performed directly on 5μl (0.1ng/μl) of the DNA extract using a 12.5μl reaction volume following the recommendations of the manufacturer (Collins et al., 2004) on a GeneAmp® PCR System 2400 (Applied Biosystems, Foster City, CA). PCR cycle conditions were set at: 95°C 11 min; 28 cycles: 94°C 1 min, 59°C 1 min, 72°C 1 min; then 60°C 60 min. The amplified products were electrophoretically separated on an ABI Prism® 3730 Genetic Analyzer (Applied Biosystems, Foster City, CA). Following separation, the fluorescent fragments were analyzed using GeneMapper-ID® version 3.1 (Applied Biosystems, Foster City, CA).

Y-STR typing

Amplification with the Y-filer™ multiplex was performed directly on 5 μl (0.1ng/μl) of the DNA extract using a 12.5 μl reaction volume following the recommendations of the

manufacturer (Mulero et al., 2006) on a GeneAmp® PCR System 2400 (Applied Biosystems, Foster City, CA). PCR conditions were: 95 °C 11 min; 30 cycles: 94°C 1 min, 61°C 1 min, 72°C 1 min; then 60°C for 80 min. The amplified products were electrophoretically separated on an ABI Prism® 3730 Genetic Analyzer (Applied Biosystems, Foster City, CA). Following separation, the fluorescent fragments were analyzed using GeneMapper-ID® version 3.1 (Applied Biosystems, Foster City, CA). Y-STR typing was carried out according to the manufacturer's protocol.

Binary polymorphism typing

Y-SNP typing was carried out with a 96 Y-SNP kit (version 2.0) manufactured by Marligen Biosciences (Ijamsville, MD). The kit comprised a core detection kit (Haplogroups A-R), a broad screening kit and eleven individual genotyping kits (Supplementary Table 1). Y-SNP fragments were amplified by multiplex PCR, where 1-10 µl (1ng/ µl) of each DNA sample was used to amplify the Y-chromosomal DNA. Amplified DNA was labelled with a fluorescent tag. The labelled products were hybridized to Y-SNP-specific oligonucleotides immobilised onto individually colour-coded beads. The Luminex 100 System (Luminex Corporation, Austin, TX) was employed to read the fluorescence of the products. Binary polymorphism typing was carried out according to the manufacturer's protocol at Marligen Biosciences (Ijamsville, MD).

Data analysis

Analysis of molecular variance (AMOVA) was carried out with Arlequin version 2.0 (Schneider et al., 2000). AMOVA measures the proportions of variance found within and between populations or groups of populations. Y-STR variation was used to calculate population pairwise R_{ST} values. Similarly, SNP variation was used to calculate pairwise Φ_{ST} values. Also, autosomal allele frequencies were used to calculate F_{ST} pairwise genetic distances. The distance matrix used consisted of the number of steps by which each pair of STR haplotypes or SNP haplogroups differed. Multidimensional scaling (MDS) plots were constructed using SPSS version 12.0.

STRUCTURE 2.1 (Pritchard et al., 2000) was used to infer clusters of individuals on the basis of the genetic data from unlinked loci alone. The model assumes the presence of K populations (where K is unknown, and is specified in advance for each run of the program). Allele frequencies at each locus characterise each assumed K. For all simulations and calculations, a no-admixture model was assumed, with allele frequencies correlated among groups. In the no-admixture model, each individual comes from one of the K populations— i.e. populations do not have mixed ancestry; anticipating that the volunteers are either of a Qahtani or Adnani ancestry. The program was run twice, once with K=1-6 as the predetermined number of Bedouin populations, and once comparing the Bedouin sample to worldwide populations also predetermined at K=1-6. Both runs used 50,000 iterations for the burn-in period and 100,000 additional iterations to obtain parameter information. Each run was carried out five times and posterior probabilities for each K were computed for each set of runs.

Median-joining networks were constructed using Network version 4.1.1.2 or 4.5.1.0 (Bandelt et al., 1999). A weighting method with a four-fold range was used in the construction of the networks, with specific weights assigned for each STR locus (Qamar et al., 2002) taking into account the Y-STR variation across the Bedouin samples. The network for haplogroup J contained many high-dimensional reticulations. This was resolved by applying the reduced-median followed by median-joining network methods.

Bayesian Analysis of Trees With Internal Node Generation—BATWING (Wilson et al., 2003)—was used to estimate population-genetic parameters and the time to most recent common ancestor (TMRCA) of a set of chromosomes. The model assumes single-step mutations of STRs and the demographic model chosen was exponential growth from an initially constant-sized population, with or without subdivision in different runs of the program. All 15 unique Y-STR loci and all variable Y-SNP markers were used. A generation time of 25 years was used as appropriate for the mutation rate estimated by Zhivotovsky *et al.*(2004), and additional parameters were as described previously (Xue et al., 2006). The 95% confidence intervals (C.I.) of posterior values for the TMRCA, effective population size ($N_{posterior}$), and expansion time were determined from the program's output.

Results

We first investigated the genetic relationships revealed by genomewide autosomal markers among the Bedouin and 14 other populations taken from the literature (Supplementary Table 2) by computing the pairwise F_{ST} genetic distances (Supplementary Table 3) from the allele frequencies of 13 STR CODIS loci (Figure 1a). Overall, with the exception of the Aniza tribe, the Bedouin clustered with the Arabic-speaking populations of the Middle East, and were distinct from Far Eastern and European populations. However, while the populations of the Persian Gulf clustered closely together, the Bedouin populations formed a much looser part of the cluster, most marked for the Aniza.

To complement this analysis, we explored the genetic clustering identified by the program STRUCTURE. There was no evidence of the presence of significant genetic structure within the Bedouin corresponding to Adnani and Qahtani clusters. The model with the highest posterior probability value was K=1 (ln $P(D)$ = −7953 for K=1 compared with −8037 for K=2; Table 1), implying that the genetic data favoured a single cluster containing all the Bedouin populations examined.

The same program was used to test for the presence of broader genetic structure when Bedouin populations were compared to a worldwide population data set (Supplementary Table 2). In the worldwide comparison, a model containing three clusters was favoured (at K=3 ln $P(D)$ = −42339; Table 1). In this analysis, almost all individuals demonstrated membership of multiple clusters, but some populations displayed membership of predominant clusters that corresponded to geographical affiliation. For example, Adima Muslim Egyptians (63%) and Coptic Egyptians (65%) showed a distinct "North African" cluster (Figure 1b, red colour). Another "European" cluster was observed among a high proportion of the South Belgians (56%, blue colour). Correspondences between clusters and populations were less clear for the the Iraqi and Qatari samples, however: individuals displayed partial membership of multiple clusters with some populations showing identities that did not correspond to geographic, linguistic or ethnic affiliation. It was noticeable that all of the Bedouin tribes with the exception of Aniza belonged predominantly to a single cluster (cluster 2, green colour). The Aniza, like the samples drawn from Qatar and Iraq, were distributed among three clusters (Supplementary Table 4).

The autosomal analyses thus indicated that the Bedouin populations resembled their geographical neighbors, but showed larger inter-population genetic distances, perhaps reflecting high levels of genetic drift, and that these tendencies were particularly marked in the Aniza.

We next analysed the Y-chromosomal lineages. Fifteen-locus Y-STR haplotypes were obtained for all individuals (n=148, the Murrah were excluded; Supplementary Table 5). Table 2 shows the Y-haplotype diversities observed in the five Bedouin samples studied,

and compared with published data from other populations. Few studies have reported 15-locus haplotypes, so it was necessary to base broader comparisons on nine loci (Supplementary Table 2). The Bedouin populations stand out even in this simple analysis because of their strikingly low diversity. With 15 loci, the Bedouin show the lowest three diversity values. With nine loci, the Bedouin population diversities were again unusually low, similar to some other Persian Gulf samples and with lower values seen only in Zriba, Berber and Somali populations.

Y-SNP analysis classified the Bedouin samples into nine haplogroups (Table 3). Haplogroup diversities were also low, ranging from zero for the Ajman—all members belonged to haplogroup J1—to 0.6 for the Aniza (Table 3). Haplogroup J1 was seen at very high frequency in the combined sample: approximately 84%—or 124 individuals—and was present at >50% in all individual populations. The other haplogroups identified were R1a1 (at 6.75%), E3b3 (6.00%), G2 (3.40%), R1b3 (1.35%) and K2, E3b1, Q* and R2 (each 0.67%).

The patrilineal genetic relationships between the populations were examined using Y-STR or Y haplogroup frequency data to calculate pairwise genetic distances (Φst) between populations, and displayed in two dimensions using MDS plots (Figure 1). The MDS plots showed the same general features as the autosomal plot (Figure 1a): clustering with their geographical neighbors, but with larger distances between the Bedouin samples. The Aniza, however, did not stand out in the Y-chromosomal analyses. We investigated the Y-chromosomal genetic structure further by carrying out an AMOVA analysis with either Y-STRs or Y-SNPs. The results were similar with the two sets of markers (Table 4), so are discussed together. When populations were grouped according to continent, a large fraction of the variation lay between groups (23% with STRs, 15% with SNPs; P<0.0001) reflecting the high degree of geographical structure also found in many other studies (Jobling et al., 2004). When the Bedouin populations were analysed alone, substantial variation was found between populations. At 18% for Y-SNPs, this exceeded the variation between continental groups— again illustrating the large genetic distances between these populations. When the Bedouin samples were grouped into Adnani and Qahtani lineages, however, no variation was found between the groups (Table 4).

To investigate the phylogenetic relationships between the Y-STR haplotypes in the Bedouin and other samples, median-joining networks were constructed. The haplogroup J1 network, containing 84% of the chromosomes is illustrated in Figure 2. It shows a high-frequency central haplotype present in most of the populations examined, both Bedouin and non-Bedouin, and a starlike structure with many radiating branches. Although individual haplotypes in the branches are often shared between populations, some clustering is evident, such as the Ajman (yellow) haplotypes in the upper right part of the network. The TMRCA — estimated using the assumption that the central haplotype is ancestral—was 3,700 years.

Modeling of the demography of the individual populations using BATWING revealed median $N_{posterior}$ estimates (see Table 5) ranging from 560 (Ajman) to 2,300 years (Shimar) and median expansion times from 2,600 (Awazim) to 8,500 years (Shimar). These expansion times are much more recent than those found using the same approach in other parts of the world (Xue et al., 2006) and point to an unusual demography for the Bedouin. The population TMRCA estimate for the Ajman sample is very recent (median 4,300 years) because of the presence of a single haplogroup, and comparable to the J1 network TMRCA. TMRCAs for the other populations reflect the coalescence of the Y lineages from different haplogroups before the formation of the Bedouin populations.

Discussion

Although previous studies have investigated classical genetic (Singh Sawhney et al., 1984; Al-Hilli, 1985; Al-Bustan et al., 2002) and molecular (Al-Nassar et al., 1981; al-Nassar et al., 1995; al-Nassar et al., 1996; Sebetan and Hajar, 1998) markers, including Y-chromosomal markers (Nebel et al., 2001) in Bedouin populations, and one Bedouin sample from Israel is present in the HGDP panel (Cann et al., 2002), the Kuwaiti Bedouin have thus far been little-studied. They share some Y-chromosomal features with the Negev Bedouins (Nebel et al., 2001): a high frequency of haplogroup J1 and low STR diversity within it, but none of the Kuwaiti individuals sampled carries the Bedouin Modal Hapolotype and thus their recent genetic history is distinct. Four main conclusions now emerge from our work:

First, the Kuwaiti Bedouin populations resemble their geographical neighbors. The autosomal markers used in this study were chosen for their ability to discriminate between individuals within any population, and this property, together with their small number, lead to high levels of noise when they are used to compare populations. Nevertheless, they reveal the genetic relationships between the Bedouin and other populations from the Arabian Peninsula (Figure 1), as expected from our understanding of the history of the region. The same conclusion is obtained with Y-chromosomal markers (Figure 1).

Second, the Bedouin exhibit evidence of small effective population sizes and substantial genetic drift. This is reflected in the low within-population diversity values for Y-chromosomal loci, and the low estimates of effective population size with BATWING; note that the latter values are estimates of the size *when growth began*—i.e. ~3-9000 years ago—and so are exceptionally low. Genetic drift leads to large distances between populations and these are seen both in the autosomal analyses—most marked for the Aniza (Figure 1a)—and the Y-chromosomal analyses (Figure 1). Again, these are exactly the genetic consequences expected from their lifestyle as small groups of desert-dwellers.

Third, no subdivision was seen according to the traditional Adnani and Qahtani lineages. This is illustrated by the lack of STRUCTURE grouping into more than one cluster (Table 1), interspersion of populations belonging to the traditional lineages in the autosomal and Y-chromosomal MDS plots (Figure 1), and lack of distinction between the lineages in the AMOVA analyses (Table 4). This finding is contrary to historical and socio-cultural expectations, and could be explained in a number of ways. For example, a common genetic pattern could have been established in an ancestral population and retained by the two groups; alternatively, there could have been sufficient gene flow between the tribes to prevent genetic differences building up; or each individual population could have diverged to such an extent that ancient relationships have been obscured. This finding is also supported by the haplotype sharing between many of the populations visualised in the network (Figure 2).

Fourth, despite not supporting the traditional binary subdivision of the Bedouin populations, the genetic data nevertheless provide considerable insights into their genetic history. The Ajman stood out in the Y-chromosomal analyses as outliers in the MDS plots (Figure 1), from their zero haplogroup and low haplotype diversity (Table 3), and from the population-specificty of their haplotypes in the network (Figure 2). A potential explanation for these findings could be the high status of the Ajman in the Bedouin social hierarchy and lack of mixing with other populations (Dickson, 1967). To this day, the Ajman tribe is known for its zealous view of consanguineous marriage and tribal relations.

The Ajman contrast with the Aniza, who show the highest diversity among the Bedouin, with six haplogroups detected (Table 3). This observation fits several aspects of their history. Their traditional territory spanned a very wide geographical area. Their first mention

in history was before the initiation of Islam where they were found in the Najd region of modern Saudi Arabia. From the start of Islam and through to the Ottoman era this tribe, via conquest and raiding, extended over all of Saudi Arabia, Syria, and North towards Iraqi Kurdistan and Jordan (Dickson, 1956). Slavery among the chieftains was a sign of power and wealth. Aniza tribes like those of Banu-Hilal and Banu Sulaym are believed to have participated in the conquest of North Africa in the 11th century CE (Abun-Nasr, 1987), possibly leading to the introduction of E-M78 chromosomes. Between the Umayyad and the Abbasid Empires, many slaves were brought from Iberia, especially Cordoba and Andalusia (Bassaam, 1977). Furthermore, envoys of the British Empire to the Persian Gulf and the Middle East between the early 19th and early 20th centuries CE, mention in their memoirs the enslaving of Georgians, Armenians and Circassians (Philby, 1923). These slaves were naturalized and given tribal affiliations under slavery abolishment treaties signed between the tribal chieftains—or emirs—and the envoys of the British Empire. This could be a reason for the presence of R1b1 and G2* Y chromosomes.

The Awazim stand out because of their high frequency of haplogroup E-M123 (24.3 %), the highest yet reported in any population. The highest frequency reported previously (23.5 %) was in Ethiopians from Amhara (Cruciani et al., 2004). The Awazim also showed the presence of haplogroup R2 (R-M124), characteristic of South Asia (Sengupta et al., 2006). There may be a possible link to the Roma ('Gypsy') migrations. It has also been proposed that the Awazim may have originated from the Caucasus—also consistent with the theory of "Suluba" (Bell and Richmond, 1937). The presence of markers M123 and M124 may also suggest that slaves brought before the initiation of Islam and after the passing of the four Caliph eras (Lewis, 1990) have been naturalized within these tribes—perhaps when slavery was abolished.

The Shimar sample carried two main haplogroups—J1 (at 52.3 %) and R1a1 (at 42.8 %)—with a small percentage of G2 (4.76 %). A historical explanation could be that the Shimar trace their origins to two regions: Iraq (Philby, 1923) and Saudi Arabia. The Shimar resided in northern Najd—currently in Saudi Arabia—after their migration during pre-Islamic times from Taye' in Yemen. Their migration to the North into Iraq ended during the 17th Century CE (Subahai, 1996; Khuraysi, 1998; Williamson and Basri, 1999). As they settled in Iraq, the Shimar gave up nomadism and became urbanized in towns and cities—most notably in Mosul (Khuraysi, 1998). It is said that the Shimar are now of two main antecedents—Iraqi Shimar and Saudi Shimar—and each refers to itself as "the Noble Tribe" (personal communication, Shimar of Kuwait). The haplogroup R1a1-M17 is found at substantial frequency in Iraq, so limited gene flow or subsequent drift (or both) could account for its presence in the Shimar— although later migrations, slavery and trade provide alternative explanations.

In all, except for the lack of subdivision into Adnani and Qahtani lineages, the genetic data fit remarkably well with the expectations for small nomadic groups in the Arabian Peninsula, and provide independent support for many aspects of the traditional histories of these populations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Abun-Nasr, JM. A History of the Maghrib in the Islamic Period. Cambridge University Press; Cambridge and New York: 1987.

Al-Bustan S, El-Zawahri M, Al-Azmi D, Al-Bashir AA. Allele frequencies and molecular genotyping of the ABO blood group system in a Kuwaiti population. Int J Hematol. 2002; 75:147–153. [PubMed: 11939260]

Al-Hilli F. The ABO and Rh blood groups in Bahrain, Arabian Gulf. Hum Biol. 1985; 57:441–447. [PubMed: 3935562]

Al-Nassar KE, Conneally PM, Palmer CG, Yu P. The genetic structure of the Kuwaiti population. I. Distribution of 17 markers with genetic distance analysis. Hum Genet. 1981; 57:192–198. [PubMed: 7228033]

al-Nassar KE, Mathew J, Thomas N, Fatania HR. HLA-DQ alpha allele and genotype frequencies in a native Kuwaiti population. Forensic Sci Int. 1995; 72:65–69. [PubMed: 7705737]

al-Nassar KE, Mathew J, Thomas N, Fatania HR. Analysis of the D1S80 (pMCT118) VNTR locus polymorphism in a native Kuwaiti population by the polymerase chain reaction. Forensic Sci Int. 1996; 78:131–138. [PubMed: 8621120]

Bandelt HJ, Forster P, Röhl A. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol. 1999; 16:37–48. [PubMed: 10331250]

Bassaam, AA. Ulamaa Najd Khilaal Sitat Quroon. Muktabat Al-Nahdha Al-Hadeethah; Damascus: 1977.

Bell, GL.; Richmond, EB. The earlier letters of Gertrude Bell. Liveright Publishing Co.; New York: 1937.

Berger B, Lindinger A, Niederstatter H, Grubwieser P, Parson W. Y-STR typing of an Austrian population sample using a 17-loci multiplex PCR assay. Int J Legal Med. 2005; 119:241–246. [PubMed: 15843993]

Cadenas AM, Zhivotovsky LA, Cavalli-Sforza LL, Underhill PA, Herrera RJ. Y-chromosome diversity characterizes the Gulf of Oman. Eur J Hum Genet. 2008; 16:374–386. [PubMed: 17928816]

Cann HM, de Toma C, Cazes L, Legrand MF, Morel V, Piouffre L, et al. A human genome diversity cell line panel. Science. 2002; 296:261–262. [PubMed: 11954565]

Carvalho M, Anjos MJ, Andrade L, Lopes V, Santos MV, Gamero JJ, et al. Y-chromosome STR haplotypes in two population samples: Azores Islands and Central Portugal. Forensic Sci Int. 2003; 134:29–35. [PubMed: 12842354]

Chang YM, Perumal R, Keat PY, Kuehn DL. Haplotype diversity of 16 Y-chromosomal STRs in three main ethnic populations (Malays, Chinese and Indians) in Malaysia. Forensic Sci Int. 2007; 167:70–76. [PubMed: 16457976]

Cherni L, Pereira L, Goios A, Loueslati BY, Khodjet el Khil H, Gomes I, et al. Y-chromosomal STR haplotypes in three ethnic groups and one cosmopolitan population from Tunisia. Forensic Sci Int. 2005; 152:95–99. [PubMed: 15939181]

Collins PJ, Hennessy LK, Leibelt CS, Roby RK, Reeder DJ, Foxall PA. Developmental validation of a single-tube amplification of the 13 CODIS STR loci, D2S1338, D19S433, and amelogenin: the AmpFlSTR® Identifiler PCR Amplification Kit. J Forensic Sci. 2004; 49:1265–1277. [PubMed: 15568700]

Cruciani F, La Fratta R, Santolamazza P, Sellitto D, Pascone R, Moral P, et al. Phylogeographic analysis of haplohroup E3b (E-M215) Y chromosomes reveals multiple migratory events within and out of Africa. Am J Hum Genet. 2004; 74:1014–1022. [PubMed: 15042509]

Dickson, HRP. Kuwait and her Neighbours. Allen & Unwin; London: 1956.

Dickson, HRP. The Arab of the Desert: a glimpse into Badawin life in Kuwait and Sau'di Arabia. Allen & Unwin; London: 1967.

Hallenberg C, Simonsen B, Sanchez J, Morling N. Y-chromosome STR haplotypes in Somalis. Forensic Sci Int. 2005; 151:317–321. [PubMed: 15939170]

Henke J, Henke L, Chatthopadhyay P, Kayser M, Dulmer M, Cleef S, et al. Application of Y-chromosomal STR haplotypes to forensic genetics. Croat Med J. 2001; 42:292–297. [PubMed: 11387642]

Jobling, MA.; Hurles, ME.; Tyler-Smith, C. Human Evolutionary Genetics. Garland Science; New York and Abingdon: 2004.

Jobling MA, Tyler-Smith C. The human Y chromosome: an evolutionary marker comes of age. Nat Rev Genet. 2003; 4:598–612. [PubMed: 12897772]

Khuraysi, H-A. Qabilat Shammar: mutaba'ah wa-tahlil. Dar al-Saqi; Beirut: 1998.

Lewis, B. Race and slavery in the Middle East: an historical enquiry. Oxford University Press; New York: 1990.

Lopez AM, Alvarez S, Gusmao L, Alves C, Mesa MS, Albentosa A, et al. Population data for 16 Y-chromosome STRs in four populations from Pyrenees (Spain). Forensic Sci Int. 2004; 140:125–129. [PubMed: 15013176]

Martinez-Jarreta B, Nievas P, Abecia E, Hinojal R, Budowle B. Haplotype distribution of nine Y-chromosome STR-loci in two northern Spanish populations (Asturias and Aragon). J Forensic Sci. 2003; 48:204–205. [PubMed: 12570231]

Mulero JJ, Chang CW, Calandro LM, Green RL, Li Y, Johnson CL, et al. Development and validation of the AmpFℓSTR® Yfiler™ PCR amplification kit: a male specific, single amplification 17 Y-STR multiplex system. J Forensic Sci. 2006; 51:64–75. [PubMed: 16423225]

Nebel A, Filon D, Brinkmann B, Majumder PP, Faerman M, Oppenheim A. The Y chromosome pool of Jews as part of the genetic landscape of the Middle East. Am J Hum Genet. 2001; 69:1095–1112. [PubMed: 11573163]

Park MJ, Lee HY, Yoo JE, Chung U, Lee SY, Shin KJ. Forensic evaluation and haplotypes of 19 Y-chromosomal STR loci in Koreans. Forensic Sci Int. 2005; 152:133–147. [PubMed: 15978339]

Philby, HSJB. The heart of Arabia: a record of travel & exploration. G. P. Putnam and Son; New York: 1923.

Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000; 155:945–959. [PubMed: 10835412]

Qamar R, Ayub Q, Mohyuddin A, Helgason A, Mazhar K, Mansoor A, et al. Y-chromosomal DNA variation in Pakistan. Am J Hum Genet. 2002; 70:1107–1124. [PubMed: 11898125]

Rebala K, Szczerkowska Z. Polish population study on Y chromosome haplotypes defined by 18 STR loci. Int J Legal Med. 2005; 119:303–305. [PubMed: 15834734]

Roewer L, Croucher PJ, Willuweit S, Lu TT, Kayser M, Lessig R, et al. Signature of recent historical events in the European Y-chromosomal STR haplotype distribution. Hum Genet. 2005; 116:279–291. [PubMed: 15660227]

Roewer L, Krawczak M, Willuweit S, Nagy M, Alves C, Amorim A, et al. Online reference database of European Y-chromosomal short tandem repeat (STR) haplotypes. Forensic Sci Int. 2001; 118:106–113. [PubMed: 11311820]

Schneider, S.; Roessli, D.; Excoffier, L. Arelquin: a software for population genetics data analysis. Genetics and Biometry Lab., Dept. of Anthropology, University of Geneva; 2000.

Sebetan IM, Hajar HA. Analysis of the short tandem repeat (STR) locus HumVWA in a Qatari population. Forensic Sci Int. 1998; 95:169–171. [PubMed: 9722978]

Sengupta S, Zhivotovsky LA, King R, Mehdi SQ, Edmonds CA, Chow CE, et al. Polarity and temporality of high-resolution Y-chromosome distributions in India identify both indigenous and exogenous expansions and reveal minor genetic influence of Central Asian pastoralists. Am J Hum Genet. 2006; 78:202–221. [PubMed: 16400607]

Singh Sawhney K, Sunderland E, Woolley V. Genetic polymorphisms in the Kuwaiti Arabs. Hum Hered. 1984; 34:303–307. [PubMed: 6479998]

Subahai, A-H-A. Qabilat Shammar wa-umara'uha. Dar Ma'd lil-Tiba'ah wa-al-Nashr wa-al-Tawzi; Dimashq: 1996.

Turrina S, Atzei R, De Leo D. Y-chromosomal STR haplotypes in a Northeast Italian population sample using 17plex loci PCR assay. Int J Legal Med. 2006; 120:56–59. [PubMed: 16328424]

Williamson, JF.; Basri, M. Qabilat Shammar al-Arabiyah, makanatuha wa-tarikhuha al-siyasi 1800-1958. Dar al-Hikmah; London: 1999.

Wilson IJ, Weale ME, Balding DJ. Inferences from DNA data: population histories, evolutionary processes and forensic match probabilities. J R Statist Soc A. 2003; 166:155–188.

Xue Y, Zerjal T, Bao W, Zhu S, Shu Q, Xu J, et al. Male demography in East Asia: a north-south contrast in human population expansion times. Genetics. 2006; 172:2431–2439. [PubMed: 16489223]

Zhivotovsky LA, Underhill PA, Cinnioglu C, Kayser M, Morar B, Kivisild T, et al. The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time. Am J Hum Genet. 2004; 74:50–61. [PubMed: 14691732]
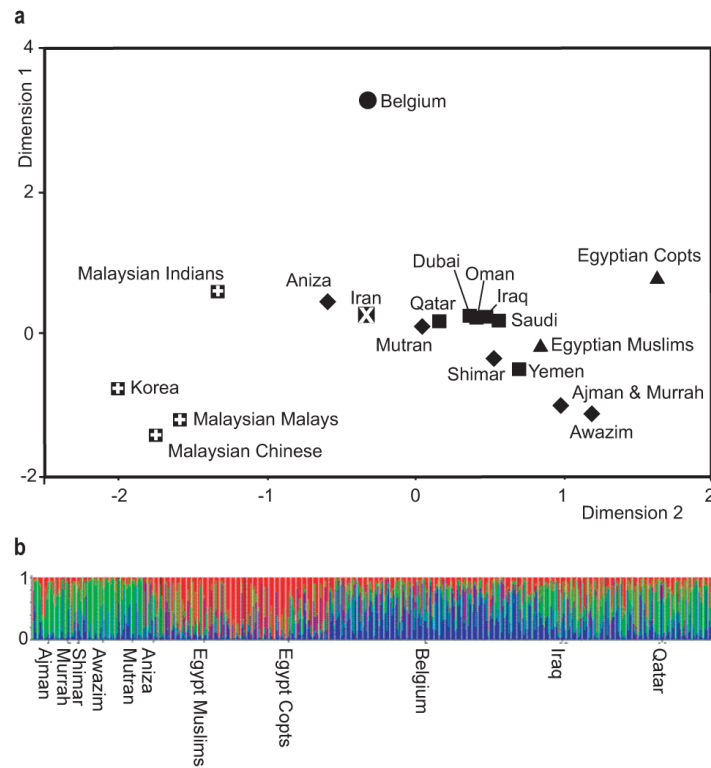
Figure 1.
Autosomal STR analyses of Kuwaiti Bedouin and additional populations. (a) MDS plot of population pairwise genetic distances (Fst values). (b) Estimated population structure based on the highest probability STRUCTURE run at K=3. Each individual is represented by a thin vertical line, which is partitioned into K coloured segments that represent the individual's estimated membership fractions in each of the K clusters. Populations are shown below the figure.
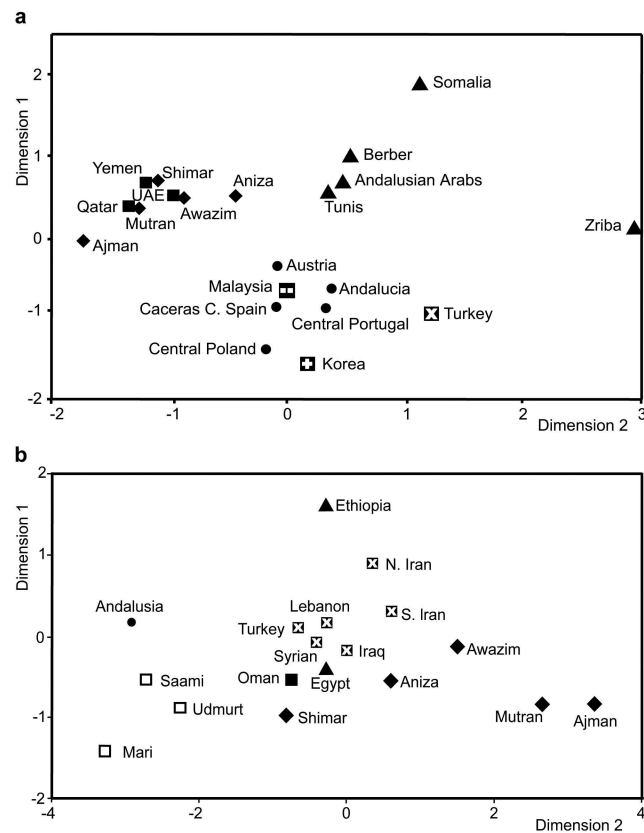
Figure 2.
Y-chromosomal analyses of Kuwaiti Bedouin and additional populations. (a) MDS plot based on population pairwise genetic distances from Y-STRs (R_{ST} values). (b) MDS plot based on population pairwise genetic distances from Y-SNPs. Population names are shown in each section.
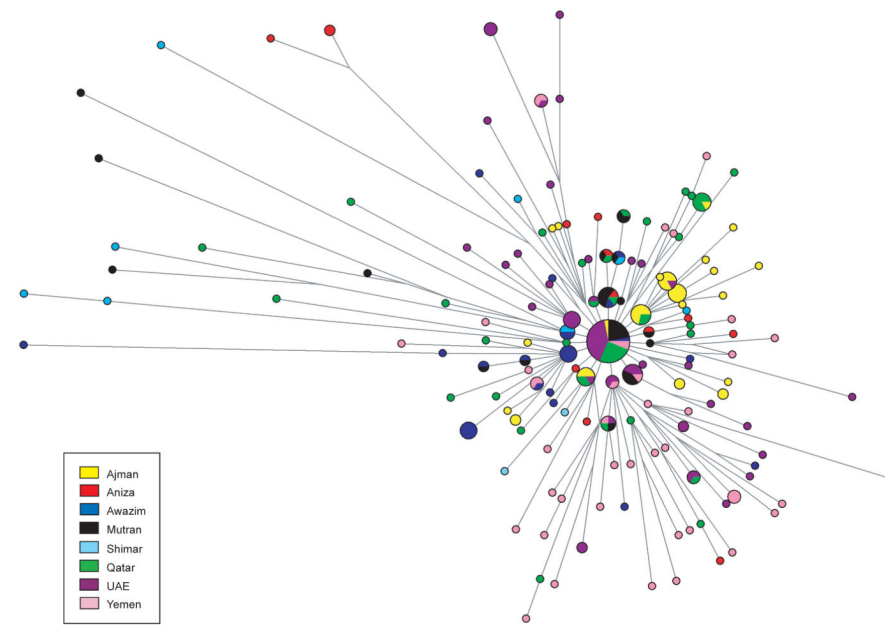
Figure 3.
Median-joining network of haplogroup J1 Y-STR haplotypes. Circles represent haplotypes
with an area proportional to frequency, and are coloured to indicate the population of origin.
Lines represent mutational differences between the haplotypes.

Table 1

STRUCTURE runs

K	ln P(D) Bedouin[a]	ln P(D) Worldwide[b]
1	−7953	−42492
2	−8037	−42401
3	−8095	−42339
4	−8029	−42560
5	−8115	−42962
6	−9782	−43433

[a]log probability of data for Adnani and Qahtani Lineages of Bedouin Populations

[b]log probability of data for Bedouin Populations in comparison to worldwide population data sets

Table 2

Haplotype diversity comparisons based on nine and 15 Y-STR loci

Group	Haplotype Diversity		Source
	15 Y-STR	**9 Y-STR**	
Aniza	0.990	0.971	Current study
Awazim	0.970	0.932	Current study
Mutran	0.920	0.598	Current study
Shimar	0.940	0.900	Current study
Ajman	0.940	0.733	Current study
Malaysian	1.000	0.998	(Chang et al., 2007)
North East Italy	1.000	-	(Turrina et al., 2006)
Tunisian Berber	1.000	0.862	(Cherni et al., 2005)
Austrian	1.000	0.995	(Berger et al., 2005)
Central Poland	1.000	0.995	(Rebala and Szczerkowska, 2005)
United Arab Emirates (UAE)	0.965	0.900	(Cadenas et al., 2008)
Qatar	0.947	0.834	(Cadenas et al., 2008)
Yemen	0.993	0.933	(Cadenas et al., 2008)
Zriba	-	0.243	(Roewer et al., 2001; Cherni et al., 2005)
Andalusian Arabs	-	0.960	(Roewer et al., 2001; Cherni et al., 2005)
Tunis	-	0.981	(Roewer et al., 2001; Cherni et al., 2005)
Andalusia	-	0.992	(Roewer et al., 2001; Martinez-Jarreta et al., 2003)
Central Portugal	-	0.991	(Roewer et al., 2001; Carvalho et al., 2003; Roewer et al., 2005)
Turkey	-	0.999	(Henke et al., 2001; Roewer et al., 2001)
Caceres Central Spain	-	0.997	(Roewer et al., 2001; Lopez et al., 2004; Roewer et al., 2005)
Somalia	-	0.919	(Roewer et al., 2001; Hallenberg et al., 2005; Roewer et al., 2005)
Korea	-	0.992	(Park et al., 2005)

Data for all nine-Y-STR comparison populations were downloaded from www.ystr.org

Table 3

Bedouin Y-SNP haplogroup data

Haplogroup	Ajman	Shimar	Aniza	Awazim	Mutran	All
			Group			
E-M123	-	-	-	24	-	9
E-M78	-	-	5	-	-	1
G-P15	-	5	14	-	3	5
J-M267	100	52	62	73	93	118
K-M70	-	-	5	-	-	1
Q-M242	-	-	-	-	3	1
R-M17	-	43	5	-	-	10
R-P25	-	-	10	-	-	2
R-M124	-	-	-	3	-	1
n	39	21	21	37	30	148
Diversity	0.00	0.56	0.60	0.40	0.12	
SD ±	0.00	0.06	0.11	0.08	0.08	

Table 4

AMOVA analysis of Y-STR and Y-SNP data in the Bedouin

Grouping	Source of Variation	Y-STRs		Y-SNPs	
		Percent Variation	Statistical Significance	Percent Variation	Statistical Significance
No grouping	Among populations	18.0	P<0.0001	22.0	P<0.0001
	Within populations	82.0	P<0.0001	78.0	P<0.0001
Adnani and Qahtani lineages	Among groups	−1.5	P=0.52	−7.6	P=0.71
	Among populations within groups	19.0	P<0.0001	27.0	P<0.0001
	Within populations	82.5	P<0.0001	80.5	P<0.0001
Continent	Among groups	23.0	P<0.0001	15.0	P<0.0001
	Among populations	6.0	P<0.0001	16.0	P<0.0001
	Within populations	71.0	P<0.0001	69.0	P<0.0001

Table 5

BATWING estimates of male demographic parameters for individual Bedouin tribes

Group	TMRCA (years)			Expansion Time (years)			N_{posterior}		
	2.5 percentile	Median	97.5 percentile	2.5 percentile	Median	97.5 percentile	2.5 percentile	Median	97.5 percentile
Ajman	2 400	4 300	8 300	1 500	3 800	14 000	250	560	1 500
Shimar	12 000	21 000	42 000	1 500	8 500	37 000	1 000	2 300	6 700
Aniza	16 000	28 000	50 000	1 200	6 000	53 000	1 300	3 000	9 000
Awazim	15 000	27 500	55 000	1 000	2 600	28 000	600	1 300	3 200
Mutran	10 500	19 000	39000	800	5 300	55 000	400	870	2 200